Compression
INFORMATION SHEET
July 2006

What is compression?

The term compression in the computer world means to make file sizes smaller by reorganizing the data within the file. Data that is duplicated or that has no value is eliminated or saved in a shorter format, greatly reducing a file's size.

Probably the most common compression format is a zip file. Files within the zip file return to their original state when unzipped and viewed.

What is imagery compression?

Imagery compression is a bit different from zipping files. Imagery compression changes the content and organization of the data within a file, not necessarily to be restored to its original condition upon opening. An imagery compression routine reorganizes the data and sacrifices data where necessary to achieve the desired level of compression. Depending on the compression ratio, this sacrifice of data may or may not be apparent to the user.

Example of image compression: large areas of the sky are the same shade of blue - only the value for one pixel needs to be saved along with the locations of the other pixels with the same color.

Why do people compress files?

Compression is done for a number of reasons:

1. Make files easier to send via email
2. Fit data on media
3. Save storage space
4. Archive data
5. Organize projects
6. Make files easier to open and use

For example, many TIFF images can be mosaicked together and compressed into one image, to make the product more usable with image viewing software than dozens of large TIFF files. Image file sizes without compression can be so large, that for the common home computer system they may become almost unusable.

Conversely, compression can make a file size so small, that the quality of the image becomes extremely poor. If the user doesn't mind poor quality, this technique is useful to save storage space or for applications such as indexing image archives.

Can imagery become better with compression?

If by "better", one means more user friendly, then yes. However, imagery cannot be made "higher quality" or "higher resolution" through compression.

So what's the downside of compression?

Files that have been compressed can lose data, even though it may not be apparent to the user. Sometimes this is not a desirable side effect, especially if done incorrectly on high quality raster data.

What's a compression ratio?

The easiest way to understand a compression ratio (shown as 50:1 or 15:1, etc.) is to think of it in terms of file size and quality. The higher the compression ratio, the smaller the file size (more data is packed into a smaller space), but the lower the quality of the compressed product.

Is prep work required to compress imagery?

Before compressing imagery, a range of image processing may be necessary to prepare images for compression, including: orthorectify or geocode the source images, grid the Digital Elevation Model (DEM) from X,Y,Z point data, mosaic and color balance the source images, combine multi-resolution imagery, perform image processing functions, add vector data.

What's Lossless and Lossy Compression?

Lossless compression uncompresses an image so its quality matches the original source - mathematically

nothing is lost. Although lossless compression sounds ideal, at times it doesn't provide enough compression and files could remain quite large.

Lossy means that the decompressed image isn't quite the same as the one you started with. Lossy compression degrades images to some degree and the more they're compressed, the more degraded they become. In many situations, such as posting images on the web or making small to medium sized prints, the image degradation isn't obvious. However, compress an image too much or enlarge an image enough, it will show.

Are there many compression formats?

Yes. However, in the geospatial arena two main players for the most part continue to dominate the market for image compression. These two companies are LizardTech, the developer of the MrSID (Multiresolution Seamless Image Database) format and ER Mapping, the developer of the ECW (Enhanced Compressed Wavelet) format. Both support JPEG 2000 compression format.

Compression file extensions:
1. MrSID MG2 and MG3 (.sid)
2. JPEG and JPEG 2000 (.jpg, .jpeg, or .jp2)
3. ER Mapping ECW (.ecw)

What formats does the Aerial Photography Field Office (APFO) use?

The USDA Farm Service Agency (FSA) and Natural Resource Conservation Service (NRCS) have used image compression for many years in order to save disk space as well as distribute and/or package data in a more convenient form.

The MrSID format is used by APFO; in large part because of a number of other government agencies were already using this format. The earlier versions of the MrSID software not only could compress the imagery but could also mosaic multiple images into a single seamless, tone-matched, compressed county mosaic. This made for an easy-to-use and pleasing to the eye base layer for GIS or hardcopy uses.

APFO has recently moved from the MrSID MG2 compression format to the MrSID MG3 compression format.
1. All Compressed County Mosaics (CCMs) under 2004 National Agriculture Imagery Program (NAIP) had the following specifications:
    a. MrSID MG2 format
    b. Compression ratio of 50:1

2. All CCMs under 2005 and 2006 NAIP have the following specifications:
    a. MrSID MG3 format
    b. Compression ratio of 15:1

Note that some compression software does more than compress data. Some compression software can reproject imagery into different coordinate systems and allow images to be added to already compressed files without having to start compression over from scratch.

Also note that for NAIP 2005 and 2006 it was decided that the superior capabilities and features of the MrSID MG3 format outweighed possible compatibility issues with ArcPad. These issues have been resolved; while ArcPad 6 does not support MG3, ArcPad 7 does. Other mobile applications using MG3 still exist by using ArcGIS or ArcView with a Pentablet or Tablet PC.

What viewers view compressed imagery?

Most image viewers and GIS software can view most compressed imagery formats.

APFO uses MG2 and MG3 for imagery.
Two free viewers that can view the MG3 format:
1. Tatuk GIS Viewer, which can be downloaded at http://www.tatukgis.com/products/viewer/viewer.aspx or via the APFO website
2. Global Mapper 6, which can be found at http://www.globalmapper.com/downloadv6.html

ArcGIS 9 can view MG3, as well as ArcGIS 8.3 with the ArcGIS 8.3 Raster Update, which can be downloaded from the ESRI website at http://support.esri.com/index.cfm?fa=downloads.patchesServicePacks.viewPatch&PID=43&MetaID=579

ArcView 3.x can view MG3 format with the appropriate plug in, which can be downloaded from LizardTech website at http://www.lizardtech.com/download/dl_download.php?detail=geo_avmrsid_dll&platform=win

All of the above software can view MG2 files as well.

Who do I contact for more information?

1. For information on compression, contact GIS Specialist, David Davis, at 801-975-3500 X278, or GIS Specialist, Brian Vanderbilt, at X240.
2. For APFO sales and product information, contact USDA-FSA-APFO at 2222 W 2300 S, Salt Lake City UT, 84119-2020, call 801-975-3503, or visit www.apfo.usda.gov.